

# *Meaning-making in online language learner interactions via desktop videoconferencing*

H. MÜGE SATAR

*Boğaziçi University, Istanbul, Turkey*  
(*email: muge.satar@boun.edu.tr*)

---

## Abstract

Online language learning and teaching in multimodal contexts has been identified as one of the key research areas in computer-aided learning (CALL) (Lamy, 2013; White, 2014).<sup>1</sup> This paper aims to explore meaning-making in online language learner interactions via desktop videoconferencing (DVC) and in doing so illustrate multimodal transcription and analysis as well as the application of theoretical frameworks from other fields. Recordings of learner DVC interactions and interviews are qualitatively analysed within a case study methodology. The analysis focuses on how semiotic resources available in DVC are used for meaning-making, drawing on semiotics, interactional sociolinguistics, nonverbal communication, multimodal interaction analysis and conversation analysis. The findings demonstrate the use of contextualization cues, five codes of the body, paralinguistic elements for emotional expression, gestures and overlapping speech in meaning-making. The paper concludes with recommendations for teachers and researchers using and investigating language learning and teaching in multimodal contexts.

Keywords: multimodal analysis, multimodal transcription, desktop videoconferencing, online language learning, semiotics, interactional sociolinguistics

---

## 1 Introduction and literature review

Recent rapid changes and improvements in telecommunication technologies have made online multimodal communication a ubiquitous part of our lives especially with increasing access to the web both on desktop and mobile devices. Most DVC tools such as *Skype* now have mobile applications that allow online multimodal communication independent of time and location. With such instant availability, the use of multimodal environments in online language learning and the effects of multimodality on online learner interactions have been identified as one of the key research areas in the field (Lamy, 2013; White, 2014).

<sup>1</sup> The data presented in this paper is based on a PhD study conducted at the Open University, UK (Satar, 2010). The theory of social presence within a community of inquiry (Rourke, Anderson, Garrison and Archer, 1999) formed the theoretical framework for the study. See Satar (2015) for details of the qualitative approach adopted for theory development, specifically for one component of the framework, i.e. sustaining interaction.

### ***1.1 Meaning-making in multimodal communication***

According to van Leeuwen (2005: 281) multimodality is a “combination of different semiotic modes – for example, language and music – in a communicative artefact or event”. Several semiotic resources can be employed including speech, writing, image, colour, layout, personal distance, movement and gaze to make “a distinctive contribution to the meaning-making process” (Sindoni, 2013: 9). Meaning-making is established through a combined and simultaneous interpretation of all available resources where the effect of each mode can only be determined through conscious reflection (Norris, 2004). In multimodal meaning-making, linguistic resources are likely to be assumed to have a dominant role. However, in this paper paralinguistic resources are not considered subordinate to language. In line with Norris (2004), it is argued that, by harnessing the power of different modes, meaning-making occurs holistically. This also resonates with Jewitt’s (2016: 70) understanding that “all modes have the potential to contribute equally to meaning”.

One further aspect of multimodal interaction emphasised by Norris (2004) is the fact that semiotic resources used by a speaker are not always interpreted by the listener in the way they were intended. She argued that meaning-making depends on the “social actors’ attention / awareness” (Norris, 2004: 151) and that is why researchers should not only analyse multimodal messages as they are transmitted, but also “how other individuals in the interaction react to these messages” (Norris, 2004: 4). When collecting, transcribing, analysing and interpreting online multimodal data, it is crucial to bear this in mind. In order to capture the full scope of the interaction, the researcher might need to obtain recordings from all interlocutors involved because depending on the internet bandwidth capacity or other technical circumstances, what is transmitted and received might not be the same.

Another challenge that researchers face is the lack of analytical frameworks specifically developed to explore language learning via online multimodal communication. In online communication, all semiotic resources “are integrated in unprecedented ways, enacting new interactional patterns and new systems of interpretation among web users” (Sindoni, 2013: 2). Therefore, it can be argued that face-to-face communication theories may not always be sufficient or appropriate when interpreting online multimodal communication.

### ***1.2 Language learner interactions in online multimodal environments***

Within the last decade, several studies have explored multimodal language learner interaction especially in synchronous video communication. In a series of studies, Wang (2004a, 2004b, 2006, 2007, 2008) looked at the nature and effects of the tutor’s use of video in online classes as well as task design and negotiation of meaning. She argued that synchronous multimodal online environments have become easier to use and are an important part of online language learning. Wang (2007) found that facial expression and gestures were used as semiotic tools for meaning-making in videoconferencing and they facilitated task completion.

A number of studies have explored language learner interactions via DVC in the context of intercultural collaborative exchanges. Most of these studies have mainly focused on the language learning potential of interaction with native speakers (Canto, Jauregi & van den Bergh, 2013; Jauregi & Banados, 2008; Lu, Goodale & Guo, 2014). However, recent research in

telecollaboration also seems to explore the multimodal features of the DVC environment. For example, Cappellini and Rivens Mompean (2015) have identified varying degrees of language learners' use of multimodal resources in teletandem exchanges.

In the context of language learner and tutor interactions via DVC, Guichon and Cohen (2014) compared videoconferencing with audioconferencing and observed more overlapping interaction in the former and more student silences in the latter. They concluded that audioconferencing did not offer paralinguistic cues for turn-taking whereas videoconferencing facilitated a rapid and seamless conversation. Stickler, Batstone, Duensing and Heins (2007) also observed longer silences in language learner–tutor interactions via audioconferencing compared to telephone conversations and postulated that lack of linguistic skills and confidence as well as availability of other semiotic modes (such as typing, raising hands and voting symbols) could have resulted in longer silences.

Lamy (2009) analysed online learner communication by adapting several methodologies including conversation analysis, affordance theory, social semiotics and geosemiotics. These combinations allowed her to better understand the multimodal nature of real-time online communication. Analysis of multimodal data necessitates a multimodal analytical approach. In this paper I will demonstrate methodologies from other fields that can be drawn on and the use of multimodal transcription and analysis methods in order to investigate meaning-making in online learner communication.

### 1.3 Multimodal transcription

Transcribing multimodal data is a complex task because the data comprises multiple modes including linguistic and paralinguistic elements, still and moving images and artefacts. Multimodal data transcription is believed to be a selective and partial process. Rapley (2007) argued that “through providing some version of a transcript you are always trying to give readers *access* to what you were able to witness” (2007: 52, original emphasis).

Some researchers believe that transcription is a prerequisite for verbal and visual data analysis as it provides initial insight thereby helping researchers become aware of salient aspects worth further exploration (Dörnyei, 2007; Swann, 2010). However, for others, especially with the advanced software available today, such as ELAN, Transana and Atlas-ti,<sup>2</sup> transcription can be seen as one of the tools that “allow the analyst to present their findings to others” (Norris, 2004: 60). For example, Develotte, Guichon and Vincent (2010), and Guichon and Cohen (2014) used ELAN to code the multimodal data directly, instead of transcribing the data first. Therefore, it might be useful for any researcher to first differentiate between transcription as an initial stage of analysis and transcription as a representation of analysis for the readers. This is an important decision to make prior to undertaking transcription as it would help determine the software or technique to be used and the level of detail needed for the transcription.

Different researchers have used different representation techniques for their transcriptions (Baldry & Thibault, 2006; Flewitt, Hampel, Hauck, Lancaster & Jewitt, 2009; Lamy &

---

<sup>2</sup> Software: Skype: <http://www.skype.com/en/>  
ooVoo: <http://www.oovoo.com/>  
Transana: <http://www.transana.org/>  
ELAN: <https://tla.mpi.nl/tools/tla-tools/elan/>  
Atlas-ti: <http://atlasti.com/>

Flewitt, 2011; Norris, 2004; Swann, 2010). For instance, Baldry and Thibault (2006) analysed advertisements in a table using a still image for each frame on the first column of the table and described the visual image, kinetic action (movement), the soundtrack and other details in the subsequent columns. Sindoni (2013) used a similar representation style in tables, but her first column included the name of the participant, followed by speech, writing, mode-switching, posture, kinetic action, gaze, staged proxemics and drawings of the participants' image. Norris (2004), however, used a number of still images representing what is visible, employed arrows or symbols to indicate movement and printed the linguistic sounds on the relevant image with different font sizes indicating emphasis. It is important to note that different techniques may suggest a dominant role for different modes; while the visual mode is the focus of Norris's (2004) method, transcription in columns may prioritise other information. For instance, the first column to the left reflects reading practices from left to right and thus, information in the first column is prioritised.

One final point to consider in multimodal transcription is the choice of appropriate transcription notations. Some analysis methods, such as Conversation Analysis, have established transcription notations like the Jefferson System (Jefferson, 2004). Multimodal analysis does not have such a universally recognised system.

#### *1.4 Multimodal analysis*

Like multimodal transcription, theories and methods for the analysis of multimodal online language learner–learner and learner–teacher interactions are still in the developmental stage within CALL. This paper draws on several theories from various fields and analysis methods including semiotics, interactional sociolinguistics, multimodal interaction analysis, theories of nonverbal communication and conversation analysis.

Semiotics studies signs and meaning-making through semiotic systems other than language (van Lier, 2004). Examples of semiotic analysis include Kress and van Leeuwen's (2001) analysis of the influence of semiotic modes on meaning-making in printed books looking at colour, layout and font. Sindoni (2013) also relied on semiotic analysis to investigate new patterns of manipulating personal distance and alternation of speech and writing in web-based videochats. Thus, semiotics provides a general theoretical framework to guide analysis of the resources employed by the participants in interaction for intentional or accidental meaning-making.

Interactional sociolinguistics (Gumperz, 1982, 2003) is a theoretical framework with its exploration of the influence of culture, background assumptions and contextualization cues on the interpretation and negotiation of meaning. According to Gumperz (1982: 131), a contextualization cue is "any feature of linguistic form that contributes to the signalling of contextual presuppositions" and which helps conversations go smoothly. One such feature Gumperz explores is the use of intonation to infer the intended meaning in discourse. In the context of multimodal interactions, paralinguistic forms can also contribute to the signalling of contextual presuppositions or assumptions that are used to infer meaning accurately.

Norris (2004) suggested that multimodal interaction analysis could be used to understand lower-level actions in multimodal interactions. These actions include gestures and body movements in the creation of social identities, relationships and practices. In analysing

online multimodal interactions, studies in nonverbal communication (Affi, 2007; Andersen, 2008; Knapp, 1980; Richmond, McCroskey & Payne, 1991) may also prove useful especially in understanding the nonverbal elements in face-to-face interactions and how these transfer to online contexts. One of these studies is Andersen's (1998, 2008) research on five codes of the body: physical appearance, kinesics (body movement), oculosics (eye behaviour), proxemics (interpersonal spatial behaviour) and haptics (tactile communication).

Although the focus of conversation analysis has been on audio recordings of face-to-face conversations, it can be argued that some of its concepts, such as overlaps, backchannels and silences in turn-taking (Jefferson, 1984; Sacks, 1992; Schegloff, 2000; Tannen, 2005, 2012) may also assist in understanding meaning-making practices online. Sacks (1992) studied turn-taking and suggested that allowing a specific amount of time between speakers, i.e. pauses or silence, ensures that only one participant speaks at a time. Overlaps or interruptions occur when more than one participant speaks at the same time. Another researcher who studied turn-taking practices was Tannen (2005, 2012). She illustrated how acceptability of overlaps and amount of silences may differ in everyday conversation according to different culturally acceptable interaction patterns. She showed that longer silences were tolerated in everyday conversations in California, whereas in New York interlocutors only tolerated a minimal pause.

Jefferson (1984) and Schegloff (2000) investigated the ways in which overlaps occur. Jefferson (1984) identified three types of overlaps: transitional, recognitional and progressional overlaps. Transitional overlaps occur when one participant takes his/her turn just before the other completes his/hers. Transitional overlaps signal enthusiastic participation. Recognitional overlaps are when the speaker attempts to anticipate and complete the unfinished sentence of another speaker. Progressional overlaps are observed when one speaker experiences disfluency and the other speaker takes the turn. On the other hand, according to Schegloff (2000), there are four types of overlaps: terminal overlaps, continuers, conditional access to the turn and chordal overlaps. Terminal overlaps are similar to transitional overlaps as identified by Jefferson (1984). Continuers are backchannels. They are the type of overlaps that index acknowledging or understanding the speaker such as "mm hm" or "uh huh". Conditional access to the turn occurs when one speaker invites the other speaker to take the turn briefly, such as when asking for help to find a word. Finally, chordal overlaps are non-serial occurrence of turns that happen at the same time, such as laughter. These are all types of non-competitive overlaps in conversation.

### ***1.5 Research questions***

With the increasing use of online multimodal communication for language learning and teaching, it is important to understand the multimodal nature of interactions and explore methodologies that are suited to investigate learners' meaning-making practices. Therefore, the guiding question for this paper is: How do language learners make meaning in their DVC interactions? In addition to investigating how semiotic resources available in Desktop Videoconferencing (DVC) shape meaning in online language learner interactions, this paper also aims to illustrate and discuss issues of multimodal transcription and analysis by providing a variety of examples.

## 2 Methods of data collection and analysis

This study followed a qualitative approach to research and used an exploratory and instrumental case study method (Creswell, 2007; Richards, 2003; Yin, 2003). Qualitative case studies permit the use of multiple sources of data and multiple analysis methods for an in-depth understanding of the phenomena being investigated.

### 2.1 Participants

The participants of the study were ten Turkish undergraduate students aged 19–22 who volunteered to participate. They were studying English Language Teaching at three different universities in different parts of Turkey. They were all in their first year of the four-year programme and were classified for the purposes of this study as advanced language learners (B2-C1). For synchronous interactions conducted via DVC, the participants were paired to constitute five cases depending on their availability for the online sessions. The data presented here are excerpts from three of the cases: Filiz and Nil, Defne and Hale, and Emre and Osman (pseudonyms). The first two cases were both female participants, while participants in the last case were both males. They all shared similar educational, linguistic and cultural backgrounds.<sup>3</sup> The participants in each pair did not know each other prior to the study.

### 2.2 Data Collection Tools

Various sources of data were collected including recordings of eighteen DVC sessions (for a total of approximately fourteen hours), interviews upon completion of the DVC sessions and questionnaires. Data from the DVC recordings were the main data analysed in this paper, while data from the interviews were used for triangulation or to provide insight into participants' individual interpretations and practices of meaning-making.

All DVC interactions were carried out in non-institutional settings, i.e. conducted outside the university, not graded and without teacher involvement. All online interactions were in English with minimal switches to Turkish, the native language of the participants. The interviews were conducted in Turkish and questionnaires were completed either in English or Turkish based on participant preferences.

The pairs took part in three or four weekly DVC sessions each lasting about an hour. Filiz and Nil completed three DVC sessions, while Defne and Hale, and Emre and Osman took part in four sessions each. In order to stimulate interpersonal interaction, the participants were provided with open-ended tasks. The first task instructed the participants to freely explore information about their interlocutor, such as details of family life, music tastes and sports. The topic of the second task was talking about personalities. The third task invited participants to talk about and compare their own rooms and an ideal room for themselves. They were then asked to describe and draw each other's rooms based on their interlocutor's description. They could draw the room either on paper or on an online whiteboard. The final task was about daily and free time activities. The participants were invited to compare their everyday and free time activities. They were encouraged to share pictures of the places and activities they were talking about.

---

<sup>3</sup> Although the participants were from Turkey, thus sharing a certain amount of cultural common ground, they lived in different parts of the country and potentially had some local cultural differences.

*ooVoo* (<http://www.oovoo.com>) was the platform used for DVC interactions. It was selected because at the time of data collection it was the only freely available DVC tool with sufficient audio and video quality that also allowed more than two interlocutors to be present simultaneously and had recording functionality. The researcher was the third participant in each session and recorded the interaction with muted sound and the camera turned off. The graphic symbol for the researcher was minimised as a small icon at the bottom right corner of the screen.

Ethical procedures were strictly followed. Approval from the ethics committee of the institution and informed consent of the participants were obtained. All participant names and any personal details used in the analysis were anonymised.

### 2.3 Data analysis techniques

The analysis of DVC recordings began by repeated viewings of the data and taking notes on the salient features of the interactions and gathering expert opinions on sections of data. In determining the salient features, social semiotics (van Lier, 2004; Kress & van Leeuwen, 2001), interactional sociolinguistics (Gumperz, 1982, 2003), multimodal interaction analysis (Norris, 2004), Andersen's (1998, 2008) five codes of the body and the concept of turn-taking in conversation analysis (Jefferson, 1984; Sacks, 1992; Schegloff, 2000; Tannen, 2005, 2012) were some of the theoretical frameworks that were drawn on (see Section 1.4). Thus, participants' meaning-making practices in DVC were explored to account for how meaning was negotiated via physical appearance, paralinguistic vocal cues, nonverbal elements that convey emotions, gestures and overlaps. Specific attention was paid to underlying shared cultural assumptions.

As discussed earlier (Section 1.3), the decision on the role of transcription is crucial in multimodal analysis. On the one hand, transcription can be an initial step for analysis by helping identify salient aspects of the data to be explored in further analysis (Dörnyei, 2007; Swann, 2010). On the other hand, multimodal transcription can be used only as a tool "to present [the] findings to others" (Norris, 2004: 60). For the present study, with fourteen hours of video data to be analysed and without a distinct framework to guide analysis, it was more feasible to embark on multimodal analysis by repeated viewings of the video data and using transcription only as a tool for representation. Therefore, all linguistic data was transcribed verbatim and, following Rapley (2007), multimodal elements in the recordings were directly annotated and coded.

Once the role of transcription was identified, it was important to choose a suitable tool for transcription and analysis. Different tools for multimodal analysis allow for different levels of detail. For example, ELAN allows the researcher to transcribe different multimodal elements in different layers which are represented simultaneously on a timeline. Based on a pilot transcription using ELAN (Figure 1), it was concluded that such transcription was better suited for researchers who have a clear theoretical framework and who use transcription as an initial stage for analysis. On the other hand, using other tools that were available, i.e. Transana and Atlas-ti, it was possible to transcribe the verbal data in a linear fashion, insert timestamps to replay the marked segments of the video data and code multimodal elements directly. Atlas-ti was selected for this study because it was possible to code not only the video data, but also all other data sources within the same software and create links amongst them. Figure 2 shows a screenshot of transcription in Atlas-ti 6. Transcription conventions are provided in Appendix.

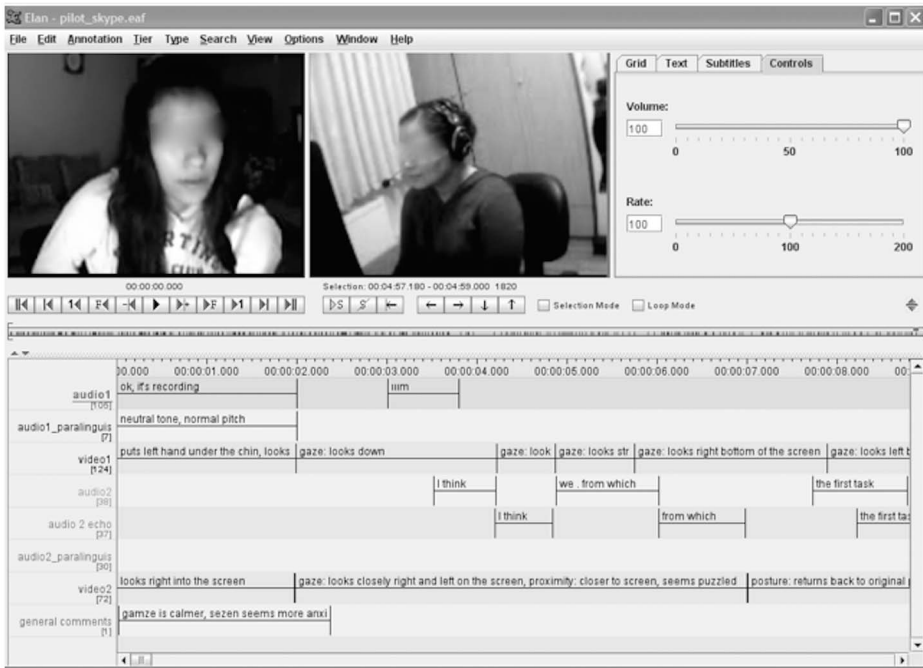


Fig. 1. Screenshot of transcription in ELAN

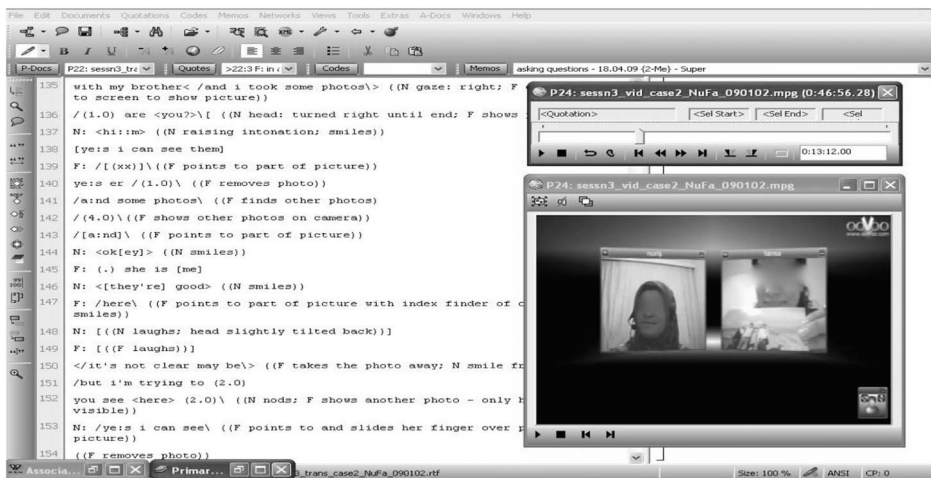


Fig. 2. Screenshot of transcription in Atlas-ti 6

### 3 Analysis

The following analysis focuses on five semiotic resources of meaning-making in language learner DVC interactions: paralinguistic contextualization cues, five codes of the body, facial expression and voice to express emotions, use of gestures and overlapping speech.



The analysis is divided into three sections. Each section investigates meaning-making practices observed in each case.

### 3.1 Paralinguistic contextualization cues and five codes of the body

The data for this section was taken from the interaction between Filiz (female) and Nil (female). Both participants were at home during their interactions. Filiz used a laptop with built-in headphones and speakers, while Nil had a desktop PC with external headphones and webcam. Nil's use of the webcam was distinctive in that she placed it to the right side of the screen and looked at the webcam instead of her screen most of the time. Nil wore a headscarf during the DVC sessions. The headscarf functioned as an artefact that marked certain interpretations of meaning as described in this section. Extract 1 below is taken from Filiz and Nil's last DVC session where both participants show each other pictures of themselves with their families.

#### Extract 1

In this extract, < and > mark start and end points for Nil's behaviour; / and \ mark that of Filiz's.

<b>Verbal</b>	<b>Nonverbal</b>
1 N: Hmm (.) here <a photo /(.)>	N laughs; F smiles
2 my sis my sister found <a photo (1.0)>\	N laughs; F smiles
3 [(xx)]	N laughs
4 F: /[oh (xx)]\	F laughs, intonation and facial expression indicates surprise and appreciation
5 N: /<(1.0)\ can you see?	N shows photo on screen - starts; N's gaze to the left indicating gaze on screen (perhaps on screen checking how well she shows the picture); F moves closer to screen
6 F: yes I can see /where is the place?\	F moves away from camera
7 /(.)\	F places her hand under the chin
8 N: name of place x	
9 F: /oh, I see>\	F smiles
10 N: <she is me>	N points to one person on the picture; laughs
11 <[and]	N starts pointing to the other person in the picture
12 F: /[yes]\	F smiles; removes her hand from the chin
13 N: another one is my sister>	N ends pointing to the other person in the picture
14 F: /<okay I see\	F moves a little away from camera
15 N: here is name of place x>	
16 <(1.0)>	N removes picture; laughs

17 F: /<also my sister is here (you) see her err>\	N gaze: right (screen); F turns head right in order to take the pictures
18 /<she is my sister (.)>\	N gaze: camera; F puts picture close to the camera
19 /<this one\	N gaze: starts looking right (screen); F points to the photo
20 N: /<yes,> I saw it>\	N: nods; F looks at the photo and points again with the other hand
21 F: she is my sister /and the other is (.) me (.)\	F points to the other person in the photo. The person is wearing a headscarf.
22 N: /<hi>\	N nods; F removes the photo
23 F: /a:nd [err]\	F shows another photo
24 N: [she is] <older than you?> is [she]	N raising intonation, gaze: right
25 F: [no she's]	
26 N: <older than you?>	N gaze: right; looks at camera when finished
27 F: /<I am older than> her\	N gaze: camera, looks right when finished; F smiling voice
28 N: <ha:>/(1.0)\	N gaze: camera; laughs; intonation surprised; F smiles
29 F: <and err	N gaze: looks right until the end
30 /this is my brother (.)\	F points to the photo
31 /and e: this is my nephew (.) [sitting here]\	F points to somebody else in the photo
32 N: [<hi:m>]	N tilts head slightly
33 F: yes err he is <my brother's son>	N nods slightly
34 (.) Muhammed /and my brother's name is\	F looks at the photo and points
35 (.) /Osman(.)\	F looks at photo
36 /<sister's name is Sultan> (.)\	N nods slightly; F points to the photo
37 N: /((smiles)) <(1.0)>\	N gaze: camera; N leans back; F removes photo; F leans back

Extract 1 starts with Nil showing a picture of her sister and herself to Filiz. The beginning of the extract (lines 1–16) is marked by laughter and smiles. Nil shows the picture (line 5) and shortly after points to the picture to show herself (line 10) and her sister (lines 11–13). In corresponding lines, Filiz moves closer to the screen (lines 5–14) to be able to see the picture better. In lines 18 and 23, Filiz shows pictures of herself, her sister and her nephew. She mirrors Nil's description and points to the people in the pictures, providing information about them. In line 21, she points to her sister in the picture who is wearing a headscarf and Nil asks whether her sister is older than Filiz. Nil finds this information surprising, which she expresses with a paralinguistic vocal cue “ha:” and a one second pause in line 28. In response, Filiz smiles and the conversation continues without further discussion with Filiz talking about the other person in the picture (lines 29–37). In line 37, both participants lean back marking closure for the topic.

It is possible to understand meaning negotiation in line 28 via the contextualization cue (ha:) coupled with an understanding of the shared social and religious culture of the participants. Filiz explained how she made sense of Nil's reaction in her interview (Extract 2).

Extract 2 (translated from Turkish)

Filiz: I was showing the photographs, ... I showed my sister, I mentioned her name and so on. Then she asked like if she was younger or older. My sister was also wearing a headscarf, I said like, I said no, when I said I am older Nil laughed there...

Interviewer: What did you think, how did you feel then?

Filiz: I wondered, well, in Turkey it's like, usually the older wears the headscarf first, then the younger one wears it. I thought maybe she found it weird for her [my sister] to be wearing the headscarf because she is younger. ... I mean I thought she thought like anyone else would do.

Filiz's comments from her interview in Extract 2 indicate that the participants were able to negotiate meaning and make lots of inferences about their partner based on the pictures they showed each other via the webcam, the paralinguistic vocal cues in the audio mode and their shared knowledge of headscarf wearing practices in society. Information gathered through these multiple modes helped index a single interpretation among many possible meanings of the paralinguistic vocal cue in line 28.

According to Andersen (1998, 2008) there are five codes of the body. The first code is physical appearance. In Extract 1, Nil's physical appearance, i.e. the fact that she was wearing a headscarf, led to certain interpretations of her actions. Moreover, in her interview, Filiz also explained that she could determine acceptable and unacceptable topics for their conversations based on Nil's video image. Filiz stated that because Nil had a headscarf and looked like a serious person, she avoided the topic of romantic relationships and did not ask her whether she had a boyfriend assuming it would not be appropriate.

In terms of kinesics, although participants had to remain in a restricted position to stay within the frame of the webcam, head nods (e.g. line 20), hand gestures (pointing to pictures, e.g. line 31) as well as forward or backward leans moving closer or away from the screen (e.g. lines 5 and 37) were semiotic resources employed for meaning-making. The head nod in line 20 reinforced what was said in the verbal mode, hand gesture in line 31 linked what is said in the verbal mode to pictures shown in the video and forward and backwards leans in lines 5 and 37 signalled interest and topic closure respectively.

It is very difficult to observe oculosics (eye behaviour) in DVC, especially when the interlocutor uses an inbuilt camera, which was the case for Filiz in Extract 1. It was relatively easy to identify Nil's gaze as she had to move her head to be able to alternate her gaze between the screen and the camera positioned next to the screen. For more information on gaze in DVC, see Satar (2013), Guichon and Cohen (2014) and Sindoni (2013).

Proxemics (interpersonal spatial behaviour) and haptics (touch) does not really exist in DVC. However, in order to show and be able to see the pictures in Extract 1, the interlocutors lean forward and backwards and bring pictures closer to the camera, creating an illusion of decreased personal distance. Similarly, it is possible to observe haptics in terms of touching and pointing to other objects, in this case to pictures.

### 3.2 *Facial expressions, voice and gestures*

Data analysed in this section was taken from the DVC interactions of Defne (female) and Hale (female). Both participants conducted the sessions in a relaxed atmosphere in their

rooms and used laptops. Both mostly looked comfortable; however, in her interview Defne mentioned her lack of practice and fluency in speaking English and instances when she struggled to understand her interlocutor. Extract 3 is one of those moments taken from their second DVC session where Defne and Hale were talking about their personality characteristics and horoscopes.






## Extract 3

<b>Verbal</b>	<b>Nonverbal</b>
1 H: Yes, we can talk about horoscopes.	
2 D: Yes.	
3 H: Are you interested in them?	
4 D: E, what? (.) sorry[	Leans forward and moves head closer to screen, reinforcing the verbal message, i.e. failure to understand
5 H: Are you interested in horoscopes?	
6 D: (4) Sorry, I could, I couldn't understand.	Unhappy facial expression. Slightly shrugs shoulders. Low tone of voice.
7 H: Are (.) you (.) interested (.) in (.) horoscopes?	Slow articulation, neutral intonation and facial expression.
8 D: Ye:s (.) okay, (.) sorry again. Err, I'm, I'm not interested in horoscopes.	Leans back, smiles, cheerful tone (suggesting relief).

In Extract 3, Hale initiates a new topic, i.e. horoscopes, and asks Defne in line 3 whether she is interested in the subject. In line 4, Defne both verbally and nonverbally indicates that she cannot understand, which triggers Hale to repeat her question. However, Defne fails to understand the question again (line 6) and her disappointment in failure to understand is clearly visible in her unhappy facial expression, low tone of voice and shrugging of shoulders. Hale repeats the question one more time with slower articulation to assist Defne. Moreover, her nonverbal behaviour is neutral without any implication of frustration due to Defne's failure to understand her. This time Defne understands the question (line 8) and her relief is expressed through her cheerful tone of voice, smiles and posture (leaning back). Failure in meaning negotiation can be face-threatening (Goffman, 1955). Drawing on semiotic resources, the participants in Extract 3 are able to express their emotions for unhappiness at failure to understand, acceptance of this failure and willingness to repeat without frustration and relief at understanding the message. Although the interaction is taking place online, the participants are observed to be socially and emotionally present as if they were face-to-face. This extract demonstrates the multimodal affordances of DVC in relaying emotions and in resolving meaning negotiation problems smoothly. The potential of DVC to transmit emotions makes it a powerful tool to meet learners' affective and social needs in online language teaching.

Extract 4 was taken from the third DVC session between Hale and Defne. The task required one participant to describe his/her dream room and the other to draw it. In this extract, Hale is describing her room and Defne is drawing it on paper when she asks where to draw the windows.

Extract 4

	<b>Verbal</b>	<b>Screenshot</b>
1	H: It's near to the bed	
2	H: from above	
3	D: Near to the bed?	
4	H: Yes, bed.	
5	D: I am bed	
6	D: and windows?	
7	H: Yes.	
8	D: Okay	

In lines 1 and 2, Hale uses her hand gestures to illustrate where to draw the windows. Likewise, in line 5, Defne uses her body to represent the bed and in line 6, she puts her hands above her head to confirm that the windows are above the bed. Hale correctly receives the nonverbal message in line 7 and Defne resumes drawing in line 8. In this extract, Hale and Defne do not use full sentences and once they negotiate meaning nonverbally, they do not focus on the language anymore. This extract is another example which shows that multimodal resources available in DVC, specifically gestures, can assist meaning negotiation in a similar way that gestures would function in face-to-face communication.

### 3.3 Overlapping speech

The data for the last analysis section was taken from the DVC interactions between Emre (male) and Osman (male). Emre joined the DVC sessions from an internet café using a desktop computer with headphones. Osman used a laptop and was at home in his room. Their interaction was marked by frequent overlaps, which is exemplified in Extract 5. The extract was taken from their second DVC session during off-task talk about the end-of-year music festivals organised at their universities. The data is analysed using theories of turn-taking behaviour (Jefferson, 1984; Schegloff, 2000; Tannen, 2005, 2012).

#### Extract 5

In this extract, underlining refers to overlapping speech and italicised words indicate the place and duration of nonverbal behaviour that co-occur with speech.

<b>Osman</b>	<b>Emre</b>	<b>Nonverbal</b>
1 We have some festivals <i>err today's</i> , nowadays in our <i>err</i> university, <i>so:</i> <i>there is</i> a big fun <i>err</i> <i>between the</i> <i>pupils here</i> (.)		<i>E: head nod</i>
2 <u>and err:</u>	<u><i>Ye:s</i></u>	<i>E: head nod</i>
3 just we are having these times, (.) <i>err</i> <i>enjoying it</i> (.) <i>err</i> and nothing else (.) so (.)		<i>E: head nod</i>
4	Yes (.)	
5 <u>Okay then err let's start our</u>	<u><i>err</i> do you know which</u> <u><i>sin..</i></u>	
6 (I)		<i>O: head nod</i> <i>implying "I am</i> <i>listening"</i>
7	Err:	
8 <u>Yes (.)</u>	<u>okay let's start</u>	
9 do you know		
10 <u>Okay, okay (xx) asking</u>	<u><i>err</i> I (need) to ask you</u>	
11	a question <i>err</i>	

12	<u>Yes, yes, yes, I'm listening</u>	<u>Do you know which</u> <u>singer err</u>	
13		coming your country, your town, your university? (1)	
14	<i>Which</i>		<i>O: head shake</i> <i>implying "I do not</i> <i>understand"</i>
15	<u>err?</u>	<u>In spring</u>	
16		festivals	
17	<i>(.) Ha:</i>	<u>Which singer?</u>	<i>O: Head nod</i>
18	yes, yes, hmm, yes, err, I, if, err, as far as I know err there are four singers		

Extract 5 starts with Osman's introduction of the topic of the end-of-year festival taking place at his university campus (lines 1–3). Emre provides verbal and nonverbal back-channels, saying yes and nodding his head to signal his attention and acknowledging Osman's turn. The verbal and nonverbal overlaps here are continuers (Schegloff, 2000).

When Osman finishes talking in line 3, he leaves a small gap with a short pause and in line 4 Emre only says yes followed by a short pause. Osman probably interprets the short pause to be a signal for the end of Emre's speech and initiates a new turn to move on with the DVC task, which overlaps with Emre's follow-up question on the topic of festivals (line 5). The overlap here is probably caused by a misalignment of the personal or cultural perception<sup>4</sup> of silence length for turn-taking (Tannen, 2005, 2012) or perhaps by the time required to construct speech in foreign language communication. In his interview, Osman expressed his lack of tolerance for silences in dyadic conversations and said that he felt "the need to continue one after another without gaps". Emre, on the other hand, stated in his interview that he needed more time to construct his sentences in English.

In line 6, Osman realises Emre's attempt to continue the off-task talk and falls silent for about a second to leave the floor to Emre, reinforcing it nonverbally with a head nod. Emre picks up the turn in line 7 with an "err:."; however, in line 8 another overlap occurs. It is possible to interpret this overlap as a progressional overlap (Jefferson, 1984). In his post-task questionnaire, Osman implied that he thought Emre's speaking skills in English were not good enough. Thus, Osman may have interpreted Emre's filler as disfluency and tried to move the conversation forward.

In lines 8 and 9, while Emre accepts Osman's earlier suggestion to continue with the task, Osman also acknowledges Emre's earlier follow-up question on the off-task topic and asks for clarification by repeating the first part of Emre's question. The misalignment of turns

<sup>4</sup> Tannen's research (2005, 2012) explored silence among interlocutors from the same country and sharing the same native language but living in two different parts of the country (i.e. California and New York). Similarly, Osman and Emre were from two separate parts of the country, i.e. north and south. Thus, the cultural differences referred to could stem from their specific cultures.

continues until line 17, when Osman understands the question and provides a response. The overlap in line 17 could be a transitional (Jefferson, 1984) or a terminal (Schegloff, 2000) overlap when Osman signals understanding of Emre's question with a contextualization cue (ha:) and nonverbal behaviour (head nod) just before Emre completes his turn.

Another possible explanation for the overlaps in lines 7–17 is potential audio/video delay in transmission. Extract 5 is a transcript of the DVC session recorded by the researcher. Thus it is impossible to determine how much audio/video delay each interlocutor experienced and how much effect such delays had on these overlaps. In his interview Emre mentioned that conversational cues were sometimes delayed, which resulted in overlapping speech. He also argued that online interactions were more difficult than face-to-face interactions, especially due to the lack or ambiguity of audio-visual conversational cues for turn-taking and the echo present in the audio channel.

The effect of the task on toleration of silences and overlaps should also be taken into consideration. The interaction in Extract 5 was taken from off-task talk, which was unstructured and spontaneous. Osman and Emre's interactions during completion of other unstructured tasks were also mostly characterised by overlapping speech. However, an exception to this was the task in their third DVC session when Osman described his dream room while Emre drew it on paper. Silences as long as 12 seconds were observed in this session as the structure of the task required one participant to describe and wait while the other drew it. Audio/visual feedback and backchannels were more useful in facilitating turn-taking in this session; Osman was able to see that Emre was busy drawing and did not feel the need to occupy the silence. Thus, although delays in transmission might be challenging at times, the semiotic resources DVC offers can facilitate turn-taking, especially when compared to voice-only online communication. Moreover, carefully structured tasks that guide learner turns would also complement efforts to overcome different cultural turn-taking practices.

#### 4 Discussion

With increased access to and use of online multimodal communication platforms in language learning and teaching, investigating multimodality to better understand learner interactions in these environments and to find appropriate research methodologies has become one of the key research areas in CALL and distance language learning and teaching (Lamy, 2013; White, 2014). The aims of this paper were to demonstrate methods of multimodal transcription and analysis and to explore the semiotic resources language learners use to make meaning in DVC interactions.

Multimodal analysis involves rich data which requires a considerable amount of time and high selectivity for transcription and analysis. Lack of established analysis frameworks and methods make it challenging to conduct research on language learning in multimodal contexts. Sections 1.3 laid out some of these challenges and showed how transcription and analysis software that meets the specific requirements of the research may help overcome some of these challenges. Section 2.3 illustrated the importance of the decision to use transcription as an initial step in analysis (Dörnyei, 2007; Swann, 2010) or as a representation of the results (Norris, 2004; Rapley, 2007). The analysis section exemplified various ways of transcription as a representation of the results. Extracts 1 and 3 used a detailed transcription of verbal and nonverbal data presented in two columns; Extract 4



included verbal data and a screenshot for nonverbal elements; whereas Extract 5 had two columns for each interlocutor's verbal output to represent overlapping speech more clearly and a third column for a description of the nonverbal output. I would argue that decisions on the role of transcription in multimodal analysis and the tools used for transcription of multimodal data are closely related to methodological choices for analysis and thus they should be well informed and carefully considered to suit the aims of the analysis.

Several theoretical frameworks were employed for the analysis of the DVC data to study meaning-making in language learner interactions. Extract 1 exemplified the use of interactional sociolinguistics (Gumperz, 1982, 2003) and Andersen's (1998, 2008) five codes of the body from nonverbal communication research. It also provided evidence of how physical appearance, contextualization cues and shared cultural background influenced meaning-making in DVC interactions. It illustrated the unique characteristics of the participants, that is, the way in which Nil's headscarf led to a certain interpretation of a paralinguistic cue based on shared cultural assumptions of scarf-wearing practices and a certain creation of identity.

Analysis of Extracts 3 and 4 drew on multimodal interaction analysis (Norris, 2004). Extract 3 illustrated how nonverbal features convey affective meaning and can, thus, express language learners' emotions such as frustration and relief related to failure in and success at meaning negotiation. Extract 4 demonstrated how learners completed the task through the use of gestures without the need to construct full sentences. This resonates with Wang's (2007) conclusion that the use of facial expressions and gestures facilitate task completion. In terms of language pedagogy, the findings indicate that DVC interactions can support learners' socio-affective communication needs and can enhance their fluency. Lu *et al.* (2014) also reported that DVC interactions positively affected learners' oral fluency. Yet as the learners can rely on semiotic resources other than language, similar to the ways they can in their face-to-face conversations, teachers or content providers should carefully plan the language tasks to trigger focus on language when the aim is to improve accuracy.

Guichon and Cohen (2014) observed more overlapping speech in videoconferencing than in audio conferencing. Similarly, frequent overlaps were observed especially in the interactions of one pair in this study. In order to investigate the nature of these overlaps, Extract 5 explored to what extent findings of turn-taking research in face-to-face settings using conversation analysis (Jefferson, 1984; Sacks, 1992; Schegloff, 2000; Tannen, 2005, 2012) could be transferred to analysis of turn-taking in DVC. These theories were partially applicable and useful in explaining the overlaps in DVC. It was relatively easy to identify continuers. Moreover, data from the interviews and post-task questionnaires suggested participants' individual or local cultural differences in conversational style for the interpretation of silences. However, delays in audio/video transmission seemed to be one of the major reasons for overlaps in DVC. In order to better understand the effects of delays on overlaps in online interaction, as Norris (2004) suggests, the conversation could be recorded as all interlocutors receive it. However, this was not possible in the current study. The requirements of the task and language learners' potential need for longer silences between turns to allow time for language production were also found to cause overlaps. Therefore, learners' awareness on the effects of audio/video delays and conversational style on turn-taking could be increased prior to interactions via DVC and learners could be advised to tolerate potential silences (Stickler *et al.* 2007) more than they would normally do in face-to-face settings.

Lamy (2009) suggested conversation analysis to be a useful approach for investigating learner interactions in online multimodal communication platforms and suggested a rearticulation of the approach drawing on affordance theory, social semiotics and geosemiotics. This paper explored the applicability of theoretical frameworks from other fields in investigating online multimodal communication among language learners. In this paper it is argued that despite certain limitations, in addition to conversation analysis, interactional sociolinguistics, theories of nonverbal communication and multimodal interaction analysis would be suitable methods in investigating meaning-making in online multimodal interactions of language learners.

## 5 Conclusion

This paper explored meaning-making in online multimodal interactions of language learners using several theories of interaction and illustrated methods of multimodal transcription and analysis. The findings and recommendations presented in this paper are limited to meaning-making in dyadic interactions by language learners who shared the same first language and the same cultural background. Further research exploring meaning-making via DVC in multi-cultural settings, i.e. in intercultural telecollaborative exchanges, would be beneficial to enhance our understanding of the role of multimodal resources in intercultural communication. Moreover, the semiotic resources that were explored here were limited to what was available in the DVC tool. Future studies may wish to investigate the role of other available semiotic resources in meaning-making in online language learning and communication contexts, such as objects present in the physical settings or the joint manipulation of online objects. Research in multimodal analysis continues to produce new tools and methods. For instance, Norris and Makboon (2015) developed “the notion of frozen actions” to investigate the use of objects in identity construction and O’Halloran (2015) reported on a new tool for multimodal analysis of video interactions, *Multimodal Analysis Video*, which has facilities for importing, viewing, transcribing and annotating videos. CALL researchers interested in exploring online multimodal language learner interactions need to follow the outcomes of research in other fields, and test the applicability and efficiency of their tools and methods to help understand multimodal online language learner interactions.

## Acknowledgements

I am grateful to the anonymous reviewers and the editors of this issue for their insightful comments and recommendations on an earlier version of this paper. I would also like to thank my PhD research supervisors, Dr Ursula Stickler and Professor James A. Coleman for their professional guidance. I am indebted to the students who kindly participated in this study.

## References

- Afifi, W. A. (2007) Nonverbal communication. In: Whaley, B. B. and Samter, W. (eds.), *Explaining communication: Contemporary theories and exemplars*. New Jersey: Erlbaum, 39–60.
- Andersen, P. A. (1998) *Nonverbal communication: Forms and functions*. Mountain View, CA: Mayfield Publishing.
- Andersen, P. A. (2008) *Nonverbal communication forms and functions* (2nd edn.). Illinois: Waveland Press Inc.

- Baldry, A. and Thibault, P. J. (2006) *Multimodal transcription and text analysis: A multimedia toolkit and coursebook*. London: Equinox.
- Canto, S., Jauregi, M. K. and van den Bergh, H. H. (2013) Integrating cross-cultural interaction through video-communication and virtual worlds in foreign language teaching programs. Burden or added value? *ReCALL*, **25**(1): 105–121.
- Cappellini, M. and Rivens Mompean, A. (2015) Role taking for teletandem pairs involved in multimodal online conversation: Some proposals for counselling practice. *Language Learning in Higher Education*, **5**(1): 243–264.
- Creswell, J. W. (2007) *Qualitative inquiry and research design: Choosing among five approaches*, (2nd edn.). London: Sage Publications.
- Develotte, C., Guichon, N. and Vincent, C. (2010) The use of the webcam for teaching a foreign language in a desktop videoconferencing environment. *ReCALL*, **22**(3): 293–312.
- Dörnyei, Z. (2007) *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Flewitt, R., Hampel, R., Hauck, M., Lancaster, L. and Jewitt, C. (2009) Multimodal data collection and transcription. In: Jewitt, C. (ed.), *The Routledge handbook of multimodal analysis*. London: Routledge, 40–53.
- Goffman, E. (1955) On face-work: An analysis of ritual elements in social interaction. *Psychiatry: Journal of Interpersonal Relations*, **18**(3): 213–231.
- Guichon, N. and Cohen, C. (2014) The impact of the webcam on an online L2 interaction. *Canadian Modern Language Journal*, **70**(3): 331–354.
- Gumperz, J. J. (1982) *Discourse strategies*. Cambridge: Cambridge University Press.
- Gumperz, J. J. (2003) Interactional sociolinguistics: A personal perspective. In: Schiffrin, D., Tannen, D. and Hamilton, H. E. (eds.), *The handbook of discourse analysis*. Oxford: Blackwell, 215–228.
- Jauregi, K. and Banados, E. (2008) Virtual interaction through video-web communication: A step towards enriching and internationalizing language learning programs. *ReCALL*, **20**(2): 183–207.
- Jefferson, G. (1984) Notes on some orderlinesses of overlap onset. In: D’Urso, V. and Leonardi, P. (eds.), *Discourse analysis and natural rhetoric*. Padua, Italy: Cleup Editore, 11–38.
- Jefferson, G. (2004) Glossary of transcript symbols with an introduction. In: Lerner, G. H. (ed.), *Conversation analysis: Studies from the first generation*. Amsterdam: John Benjamins, 13–31.
- Jewitt, C. (2016) Multimodal analysis. In: Georgakopoulou, A. and Spilioti, T. (eds.), *The Routledge handbook of language and digital communication*. Abingdon and New York: Routledge, 69–84.
- Knapp, M. L. (1980) *Essentials of nonverbal communication*. New York: Holt, Rinehart and Winston.
- Kress, G. R. and van Leeuwen, T. (2001) *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Lamy, M. N. (2009) Multimodality in second language conversations online: Looking for a methodology. In: Baldry, A. and Montagna, E. (eds.), *Interdisciplinary perspectives on multimodality: Theory and practice*. Campobasso, Italy: Palladino, 385–403. <https://edutice.archives-ouvertes.fr/edutice-00387133/document>
- Lamy, M. N. (2013) Distance CALL online. In: Thomas, M., Reinders, H. and Warschauer, M. (eds.), *Contemporary computer-assisted language learning*. London: Bloomsbury, 141–158.
- Lamy, M. N. and Flewitt, R. (2011) Describing online conversations: Insights from a multi-modal approach. In: Develotte, C., Kern, R. and Lamy, M. N. (eds.), *Décrire la conversation en ligne: Le face à face distanciel*. [Describing online conversation: face-to-face in distance mode]. Lyon, France: ENS Editions, 71–94.
- Lu, R., Goodale, T. and Guo, Y. (2014) Impact of videoconference with native English speakers on Chinese EFL learners’ oral competence and self-confidence. *Open Journal of Social Sciences*, **2**(2): 54–60.
- Norris, S. (2004) *Analyzing multimodal interaction: A methodological framework*. New York: Routledge.

- Norris, S. and Makboon, B. (2015) Objects, frozen actions, and identity: A multimodal (inter)action analysis. *Multimodal Communication*, **4**(1): 43–59.
- O'Halloran, K. L. (2015) Multimodal digital humanities. In: Pericles Trifonas, P. (ed.), *International handbook of semiotics*. London: Springer, 389–415.
- Rapley, T. (2007) *Doing conversation, discourse and document analysis*. London: Sage Publications.
- Richards, K. (2003) *Qualitative inquiry in TESOL*. Basingstoke: Palgrave Macmillan.
- Richmond, V. P., McCroskey, J. C. and Payne, S. K. (1991) *Nonverbal behavior in interpersonal relations* (2nd edn.). New Jersey: Prentice Hall.
- Rourke, L., Anderson, T., Garrison, D. R. and Archer, W. (1999) Assessing social presence in asynchronous text-based computer conferencing. *The Journal of Distance Education*, **14**(2): 50–71.
- Sacks, H. (1992) *Lectures on conversation*. Cambridge, MA: Blackwell.
- Satar, H. M. (2010) *Social presence in online multimodal communication: A framework to analyse online interactions between language learners* (Unpublished PhD thesis). The Open University, Milton Keynes, UK.
- Satar, H. M. (2013) Multimodal language learner interactions via desktop videoconferencing within a framework of social presence: Gaze. *ReCALL*, **25**(1): 122–142.
- Satar, H. M. (2015) Sustaining multimodal language learner interactions online. *CALICO Journal*, **32**(3): 480–507.
- Schegloff, E. (2000) Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, **29**(1): 1–63.
- Sindoni, M. G. (2013) *Spoken and written discourse in online interactions: A multimodal approach*. London: Routledge.
- Stickler, U., Batstone, C., Duensing, A. and Heins, B. (2007) Distant classmates: Speech and silence in online and telephone language tutorials. *European Journal of Open, Distance and E-Learning*, **2007**(2). [http://www.eurodl.org/materials/contrib/2007/Stickler\\_Batstone\\_Duensing\\_Heins.html](http://www.eurodl.org/materials/contrib/2007/Stickler_Batstone_Duensing_Heins.html)
- Swann, J. (2010) Transcribing spoken interaction. In: Hunston, S. and Oakey, D. (eds.), *Introducing applied linguistics: Concepts and skills*. London: Routledge, 163–176.
- Tannen, D. (2005) *Conversational style: Analyzing talk among friends*. New York: Oxford University Press.
- Tannen, D. (2012) Turn-taking and intercultural discourse and communication. In: Paulston, C., Kiesling, S. and Rangel, E. (eds.), *The handbook of intercultural discourse and communication*. Chichester: Wiley-Blackwell, 135–157.
- van Leeuwen, T. (2005) *Introducing social semiotics*. Abingdon: Routledge.
- van Lier, L. (2004) *The ecology and semiotics of language learning: A sociocultural perspective*. Norwell, MA: Kluwer Academic Publishers.
- Wang, Y. (2004a) Supporting synchronous distance language learning with desktop videoconferencing. *Language Learning & Technology*, **8**(3): 90–121.
- Wang, Y. (2004b) Distance language learning: Interactivity and fourth-generation internet-based videoconferencing. *CALICO Journal*, **21**(2): 373–395.
- Wang, Y. (2006) Negotiation of meaning in desktop videoconferencing-supported distance language learning. *ReCALL*, **18**(1): 122–145.
- Wang, Y. (2007) Task design in videoconferencing-supported distance language learning. *CALICO Journal*, **24**(3): 591–630.
- Wang, Y. (2008) *Distance language learning and desktop videoconferencing*. Saarbrücken, Germany: VDM Verlag Dr. Müller.
- White, C. (2014) The distance learning of foreign languages: A research agenda. *Language Teaching*, **47**(4): 538–553.
- Yin, R. K. (2003) *Case Study Research: Design and Methods* (3rd edn.). Thousand Oaks: Sage Publications.

## Appendix

### *Transcription conventions*

Symbol	Name	Use
[ text ]	Brackets	Indicates the start and end points of overlapping speech
(# of seconds)	Timed Pause	A number in parentheses indicates the time, in seconds, of a pause in speech
(.)	Micropause	A brief pause, usually less than 0.2 seconds
.	Period	Indicates falling pitch or intonation
?	Question Mark	Indicates rising pitch or intonation
,	Comma	Indicates a temporary rise or fall in intonation
:::	Colon(s)	Indicates prolongation of a sound.
(xx)	Parentheses	Speech which is unclear or in doubt in the transcript.
((text))	Double Parentheses	Annotation of non-verbal activity.

Adapted from Jefferson (1984). Notations specific to Extracts 1 and 5 are stated at the beginning of these extracts.